



## Human / AI interaction loop training as a new approach for interactive learning with reinforcement-learning agents

Neda Navidi

AI-R Inc., 400 McGill st., Montreal, Canada

### Abstract:

Human / AI interaction loop training as a new approach for interactive learning with reinforcement-learning: Reinforcement-Learning (RL) in various decision-making tasks of Machine-Learning (ML) provides effective results with an agent learning from a stand-alone reward function. However, it presents unique challenges with large amounts of environment states and action spaces, as well as in the determination of rewards. This complexity, coming from high dimensionality and continuousness of the environments considered herein, calls for a large number of learning trials to learn about the environment through RL. Imitation-Learning (IL) offers a promising solution for those challenges, using a teacher's feedback. In IL, the learning process can take advantage of human-sourced assistance and/or control over the agent and environment. In this study, we considered a human teacher, and an agent learner. The teacher takes part in the agent's training towards dealing with the environment, tackling a specific objective, and achieving a predefined goal. Within that paradigm, however, existing IL approaches have the drawback of expecting extensive demonstration information in long-horizon problems. With this work, we propose a novel approach combining IL with different types of RL methods, namely State-action-reward-state-action (SARSA) and Proximal Policy Optimization (PPO), to take advantage of both IL and RL methods. We address how to effectively leverage the teacher's feedback - be it direct binary or indirect detailed - for the agent learner to learn sequential decision-making policies. The results of this study on various OpenAI-Gym environments show that this algorithmic method can be incorporated with different RL-IL combinations at different respective levels, leading to significant reductions in both teacher effort and exploration costs.

### Biography:

Neda has completed her PhD in autonomous driving field from École de Technologie Supérieure (ÉTS), and



postdoctoral studies from HEC Montréal, McGill University and Polytechnique Montréal. She has been machine learning (ML) researcher, applied research scientist and data scientist in different research teams. She is also an expert in deep learning, reinforcement learning, supervised / unsupervised learning, natural language processing, computer vision, and time series data. She now works in AI research and development at AI Redefined Inc.

### Publication of speakers:

1. Neda Navidi, Developing Equations to Predict Waist Circumference Measurements based on the National Heart, Lung, and Blood Institute Method from the World Health Organization Method, *Annals of Epidemiology*, Available online 22 August 2020, In Press, Journal Pre-proof
2. Takahashi Hiroshi , Masahiro Nakajima , Kouichi Ozaki , Toshihiro Tanaka , Naoyuki Kamatani1 , Shiro Ikegawa (2010).: Prediction model for knee osteoarthritis based on genetic and clinical information. *Arthritis Research & Therapy* 12:R187.
3. Teitel A.D, Zieve D. MedlinePlus Medical Encyclopedia. National Institutes of Health. "Osteoarthritis." Last updated: Sept 26, 2013. <http://www.nlm.nih.gov/medlineplus/ency/article/000423.htm>
4. World Health Organization. "Chronic Rheumatic Conditions." Chronic diseases and health promotion. 2014. <http://www.who.int/chp/topics/rheumatic/en/>

[International Conference on Ethicalhacking and Cybersecurity | May 21, 2020 | London, UK](#)

**Citation:** Neda Navidi; Human / AI interaction loop training as a new approach for interactive learning with reinforcement-learning agents; *Cyber Security* 2020; May 21, 2020; London, UK